Document made available under the Patent Cooperation Treaty (PCT)

International application number: PCT/US05/000053

International filing date: 03 January 2005 (03.01.2005)

Document type: Certified copy of priority document

Document details: Country/Office: US
Number: 60/533,894
Filing date: 02 January 2004 (02.01.2004)

Date of receipt at the International Bureau: 29 August 2005 (29.08.2005)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b)

1358551

THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE

United States Patent and Trademark Office

August 18, 2005

THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY FROM THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK OFFICE OF THOSE PAPERS OF THE BELOW IDENTIFIED PATENT APPLICATION THAT MET THE REQUIREMENTS TO BE GRANTED A FILING DATE.

APPLICATION NUMBER: ***60/533,894***
FILING DATE: ***January 02, 2004***
RELATED PCT APPLICATION NUMBER: ***PCT/US05/00053***

UNITED STATES PATENT AND TRADEMARK OFFICE
DEPARTMENT OF COMMERCE

Certified by

Jon W. Dudas

Under Secretary of Commerce
for Intellectual Property
and Director of the United States
Patent and Trademark Office

16523 U.S. PTO
010204

PROVISIONAL APPLICATION FOR PATENT COVER SHEET

This is a request for filing a PROVISIONAL APPLICATION FOR PATENT under 37 CFR 1.53(c).

Express Mail Label No. EL984956530US

INVENTOR(S)

Given Name (first and middle [if any])	Family Name or Surname	Residence (City and either State or Foreign Country)
Benjamin L.	Miller	17 Rolling Meadows Way Penfield, NY 14526
Christopher M.	Strohsahl	P.O. Box 353 Saugerties, New York 12477

☐ *Additional inventors are being named on the ____ separately numbered sheets attached hereto*

16591 U.S. PTO
60/533894
010204

TITLE OF THE INVENTION (280 characters max)

METHOD OF IDENTIFYING HAIRPIN DNA PROBES BY PARTIAL FOLD ANALYSIS

Direct all correspondence to: **CORRESPONDENCE ADDRESS**

☐ Customer Number [ ] → *Place Customer Number Bar Code Label here*

OR *Type Customer Number here*

☒ Firm *or* Individual Name	Edwin V. Merkel				
Address	Nixon Peabody LLP				
Address	Clinton Square, P.O. Box 31051				
City	Rochester	State	NY	ZIP	14603-1051
Country	USA	Telephone	(585) 263-1128	Fax	(585) 263-1600

ENCLOSED APPLICATION PARTS ***(check all that apply)***

☒ Specification *Number of Pages* 8 ☐ CD(s), Number

☐ Drawing(s) *Number of Sheets* ☐ Other (specify)

☐ Application Data Sheet. See 37 CFR 1.76

METHOD OF PAYMENT OF FILING FEES FOR THIS PROVISIONAL APPLICATION FOR PATENT

☒ Applicant claims small entity status. See 37 CFR 1.27.

☒ A check or money order is enclosed to cover the filing fees

☒ The Commissioner is hereby authorized to charge fees which may be required, or credit any overpayment to Deposit Account Number: 14-1138

FILING FEE AMOUNT ($): $80.00

☐ Payment by credit card. Form PTO-2038 is attached.

The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government.

☒ No.

☐ Yes, the name of the U.S. Government agency and the Government contract number are: ____________________

Respectfully submitted,

SIGNATURE [signature]

Date 1/2/2004

REGISTRATION NO. *(if appropriate)* 40,087

TYPED or PRINTED NAME Edwin V. Merkel

Docket Number: 176/61750 (1269)

TELEPHONE (585) 263-1128

USE ONLY FOR FILING A PROVISIONAL APPLICATION FOR PATENT

Mail Stop Provisional Patent Application
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

R738472.1

FEE TRANSMITTAL FOR FY 2003

Patent fees are subject to annual revision.

☒ Applicant claims small entity status. See 37 CFR 1.27

TOTAL AMOUNT OF PAYMENT ($) 80

Complete if Known	
Application Number	
Filing Date	
First Named Inventor	Miller et al.
Examiner Name	
Art Unit	
Attorney Docket No.	176/61750 (1269)

METHOD OF PAYMENT *(check all that apply)*

☒ Check ☐ Credit Card ☐ Money Order ☐ Other ☐ None

☐ Deposit Account:

Deposit Account Number: 14-1138

Deposit Account Name: Nixon Peabody LLP

The Commissioner is authorized to: *(check all that apply)*

☐ Charge fee(s) indicated below ☒ Credit any overpayments

☒ Charge any additional fee(s)

☐ Charge fee(s) indicated below, except for the filing fee to the above-identified deposit account.

FEE CALCULATION

1. BASIC FILING FEE

Large Entity Fee Code	Large Entity Fee ($)	Small Entity Fee Code	Small Entity Fee ($)	Fee Description	Fee Paid
1001	770	2001	385	Utility filing fee	
1002	340	2002	170	Design filing fee	
1003	530	2003	265	Plant filing fee	
1004	770	2004	385	Reissue filing fee	
1005	160	2005	80	Provisional filing fee	80

SUBTOTAL (1) ($) 80

2. EXTRA CLAIM FEES FOR UTILITY AND REISSUE

			Extra Claims		Fee from below		Fee Paid
Total Claims		-20** =		X		=	0
Independent Claims		-3** =		X		=	0
Multiple Dependent				X		=	0

Large Entity Fee Code	Large Entity Fee ($)	Small Entity Fee Code	Small Entity Fee ($)	Fee Description
1202	18	2202	9	Claims in excess of 20
1201	86	2201	43	Independent claims in excess of 3
1203	290	2203	145	Multiple dependent claim, if not paid
1204	86	2204	43	** Reissue independent claims over original patent
1205	18	2205	9	** Reissue claims in excess of 20 and over original patent

SUBTOTAL (2) ($) 0

**or number previously paid, if greater; For Reissues, see above

FEE CALCULATION (continued)

3. ADDITIONAL FEES

Large Entity Fee Code	Large Entity Fee ($)	Small Entity Fee Code	Small Entity Fee ($)	Fee Description	Fee Paid
1051	130	2051	65	Surcharge – late filing fee or oath	
1052	50	2052	25	Surcharge - late provisional filing fee or cover sheet	
1053	130	1053	130	Non-English specification	
1812	2,520	1812	2,520	For filing a request for *ex parte* reexamination	
1804	920*	1804	920*	Requesting publication of SIR prior to Examiner action	
1805	1,840*	1805	1,840*	Requesting publication of SIR after Examiner action	
1251	110	2251	55	Extension for reply within first month	
1252	420	2252	210	Extension for reply within second month	
1253	950	2253	475	Extension for reply within third month	
1254	1,480	2254	740	Extension for reply within fourth month	
1255	2,010	2255	1,005	Extension for reply within fifth month	
1401	330	2401	165	Notice of Appeal	
1402	330	2402	165	Filing a brief in support of an appeal	
1403	290	2403	145	Request for oral hearing	
1451	1,510	1451	1,510	Petition to institute a public use proceeding	
1452	110	2452	55	Petition to revive – unavoidable	
1453	1,330	2453	665	Petition to revive – unintentional	
1501	1,330	2501	665	Utility issue fee (or reissue)	
1502	480	2502	240	Design issue fee	
1503	640	2503	320	Plant issue fee	
1460	130	1460	130	Petitions to the Commissioner	
1807	50	1807	50	Processing fee under 37 CFR 1.17(q)	
1806	180	1806	180	Submission of Information Disclosure Stmt	
8021	40	8021	40	Recording each patent assignment per property (times number of properties)	
1809	770	2809	385	Filing a submission after final rejection (37 CFR 1.129(a))	
1810	770	2810	385	For each additional invention to be examined (37 CFR 1.129(b))	
1801	770	2801	385	Request for Continued Examination (RCE)	
1802	900	1802	900	Request for expedited examination of a design application	

Other fee (specify) ______________

*Reduced by Basic Filing Fee Paid

SUBTOTAL (3) ($) 0

CERTIFICATE OF MAILING OR TRANSMISSION [37 CFR 1.8(a)]

I hereby certify that this correspondence is being:

☐ deposited with the United States Postal Service on the date shown below with sufficient postage as first class mail in an envelope addressed to: Mail Stop ________, Commissioner for Patents, P. O. Box 1450, Alexandria, VA 22313-1450

☐ transmitted by facsimile on the date shown below to the United States Patent and Trademark Office at (703) ____________.

Date ______________ Signature ______________

Typed or printed name ______________

SUBMITTED BY				Complete *(if applicable)*	
Name *(Print/Type)*	Edwin V. Merkel	Registration No. *(Attorney/Agent)*	40,087	Telephone	(585) 263-1128
Signature	Edwin V. Merkel			Date	January 2, 2004

SEND TO: Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

R738477.1

EXPRESS MAIL CERTIFICATE

DOCKET NO.: **176/61750 (1269)**

APPLICANTS: **Benjamin L. Miller and Christopher M. Strohsahl**

TITLE: **METHOD OF IDENTIFYING HAIRPIN DNA PROBES BY PARTIAL FOLD ANALYSIS**

Certificate is attached to the **Provisional Application for Patent Cover Sheet (1 page) and Fee Transmittal (1 page)** of the above-identified application.

"EXPRESS MAIL" NUMBER: **EL984956530US**
DATE OF DEPOSIT: **January 2, 2004**

I hereby certify that this paper or fee is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR 1.10 on the date indicated above and is addressed to Mail Stop Provisional Patent Application, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.

Edwin V. Merkel
(Typed or Printed Name of Person Mailing Paper or Fee)

(Signature of Person Mailing Paper or Fee)

EXPRESS MAIL CERTIFICATE

DOCKET NO.: **176/61750 (1269)**

APPLICANTS: **Benjamin L. Miller and Christopher M. Strohsahl**

TITLE: **METHOD OF IDENTIFYING HAIRPIN DNA PROBES BY PARTIAL FOLD ANALYSIS**

Certificate is attached to the **Provisional Patent Application (8 pages)** of the above-identified application.

"EXPRESS MAIL" NUMBER: **EL984956530US**
DATE OF DEPOSIT: **January 2, 2004**

I hereby certify that this paper or fee is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR 1.10 on the date indicated above and is addressed to Mail Stop Provisional Patent Application, Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.

Edwin V. Merkel
(Typed or Printed Name of Person Mailing Paper or Fee)

Ed Merkel

(Signature of Person Mailing Paper or Fee)

TITLE: METHOD OF IDENTIFYING HAIRPIN DNA PROBES BY PARTIAL FOLD ANALYSIS

INVENTORS: BENJAMIN L. MILLER and CHRISTOPHER M. STROHSAHL

DOCKET NO: 176/61750 (UR 1269)

U.S. PROVISIONAL PATENT APPLICATION

METHOD OF IDENTIFYING HAIRPIN DNA PROBES BY PARTIAL FOLD ANALYSIS

BACKGROUND OF THE INVENTION

The use of DNA hairpins as molecular beacons, both in solution (Broude, *Trends Biotechnol.* 20:249–256 (2002); Dubertret et al., *Nat. Biotehcnol.* 19:365-370 (2001)) and immobilized on a solid surface (Fang et al., *J. Am. Chem. Soc.* 121:2921-2922 (1999); Wang et al., *Nucl. Acids. Res.* 30:e61 (2002); Du et al., *J. Am. Chem. Soc.* 125:4012–4013 (2003)), has proven to be an excellent method for "label–free" detection (Chan et al., *J. Am. Chem. Soc.* 123:11797-11798 (2001)) of biological entities. This disclosure describes a new method of molecular beacon discovery which relies on the generation of naturally occurring hairpins. The method of discovery and its advantages shall be discussed herein.

The traditional method of molecular beacon generation is to supplement a naturally occurring DNA sequence at both the 5' and 3' ends with the necessary nucleotide composition to force the formation of a hairpin. This technique has a major flaw in that the introduction of nucleotides that are not specific for the intended target sequence increases the likelihood of non-specific binding. The use of naturally occurring DNA hairpins obviates this flaw by eliminating the need for supplementation of additional bases, the result: a probe that is completely specific for its designed target.

DESCRIPTION OF THE INVENTION

The method of the invention involves obtaining or providing a nucleotide sequence from a molecular target. The nucleotide sequence can be sequenced from an isolated cDNA or obtained from an online database such as GenBank. Regardless of the source of the nucleotide sequence, a partial fold analysis is performed on the nucleotide sequence using any of a variety of suitable folding software such as, e.g., RNAStructure program (available from D. Turner at the University of Rochester, Rochester, NY), Mfold software package (available from M. Zucker at the Rensselear Polytechnic Institute, Rensselear, NY), and Vienna RNA software package, including RNAfold, RNAeval, and RNAsubopt (available from I. Hofacker at the Institute for Theoretical Chemistry, Vienna Austria). The resulting folded structure may or may not be the true active conformation of the RNA molecule in a cellular environment; however, it represents the lowest free energy state as predicted using such software. It is believed that more often than not, the predicted lowest free energy state of the nucleic acid molcecule sufficiently resembles the true active conformation. Nonetheless, the resulting folded structure is analyzed to identify hairpin regions thereof.

Having identified hairpin structures within the folded structure of the prospective target nucleic acid molecule, the hairpin sequences are isolated from the larger sequence (i.e., that was used as input to the folding software). The isolation can be performed *in silico*. Once isolated, the hairpin sequence is subjected to a second structural prediction as was performed on the prospective target nucleic acid molecule.

The overall length of the selected hairpin is preferably between about 12 and about 60 nucleotides, more preferably between about 20 and about 50 nucleotides, most preferably between about 30 and about 40 nucleotides. It should be appreciated, however, that longer or shorter nucleic acids can certainly be used. According to the preferred hairpins, the regions forming the stem of the hairpin are preferably at least about 4

nucleotides in length and up to about 28 nucleotides in length, depending on the overall length of the nucleic acid probe and the size of a loop region present between the portions forming the stem. It is believed that a loop region of at least about 4 or 5 nucleotides is need to form a stable hairpin. The regions forming the stem can be perfectly matched (i.e., having 100 percent complementary sequences that form a perfect stem structure of the hairpin conformation) or less than perfectly matched (i.e., having non-complementary portions that form bulges within a non-perfect stem structure of the hairpin conformation). When the first and second regions are not perfectly matched, the regions forming the stem structure can be the same length or they can be different in length.

Importantly, applicants have found that the predicted E value for the hairpin should preferably be at most about -3 kcal/mol, more preferably at most about -3.5 kcal/mol, most preferably between about -4 kcal/mol and about -12 kcal/mol. It should be appreciated, however, that identified hairpins can still function as molecular probes if their predicted E value falls outside these ranges.

Once the structure of the hairpin itself has been predicted, the duplex formed between the hairpin and its complement is subjected to a structural prediction as was performed on the prospective target nucleic acid molecule and the hairpin. This step, not necessary for identification of the hairpin *per se*, is performed primarily to ensure that the hybridization of the two sequences (hairpin and complement), and thus the disruption of the hairpin, will be an energetically favorable process. Ideally, there should be an increase in the predicted E value, preferably at least about a two-fold increase, preferably at least about a five-fold increase, more preferably at least about a ten-fold increase. This structural prediction also serves to demonstrate the primary advantage of the technique: after hybridization, there are no extraneous unhybridized nucleotides and, thus, lowered risk of non–specific binding.

To further verify the specificity of the hairpin sequence for its complement, the hairpin sequence can be used to perform a BLAST database search (of, e.g., the GenBank database). Ideally, the resulting BLAST search will show not only high match scores for molecular targets (or target organisms), but also a sharp discrepancy (or clear demarcation) between the high match scores of the target and any match scores of nucleic acid molecules bearing lower similarity. By sharp discrepancy and clear demarcation, it is intended that a gap of at least about 5 points, preferably at least about 10 points, more preferably at least about 15 points, most preferably at least about 20 points, exists between the target and non-target sequences. This is exemplified in Example 1 below.

The probes identified in accordance with the present invention can be used in any of a variety of hybridization-based applications, typically though not exclusively detection procedures for identifying the presence in a sample of a target nucleic acid molecule. By way of example, uses of the probes are described in greater detail in U.S. Utility Patent Application to Miller et al., entitled "Hybridization-Based Biosensor Containing Hairpin Probes and Use Thereof," filed concurrently with this application and expressly incorporated by reference in its entirety.

<u>Example 1</u> - Hairpins Targeted to *Bacillus anthracis pag* Gene

A partial gene sequence of the *Bacillus anthracis* Pag gene (isolate IT – Carb3 – 6254) (Adone et al., *J. Appl. Microbiol.* 92:1-5 (2002), which is hereby incorporated by reference in its entirety) was obtained from GenBank. The secondary structure of ~1000 nucleotide fragments of the aforementioned sequence were then computationally predicted (RNAstructure v. 3.7: Mathews et al., *J. Mol. Biol.* 288:911–940 (1999), which is hereby

incorporated by reference in its entirety). Ideally, the secondary structure of the entire sequence would be predicted, but it was discovered repeatedly that segments larger than approximately 1000 bases would crash the program RNAstructure v. 3.7.

An example of a large sequence structure prediction is shown in Figure 1 (below).

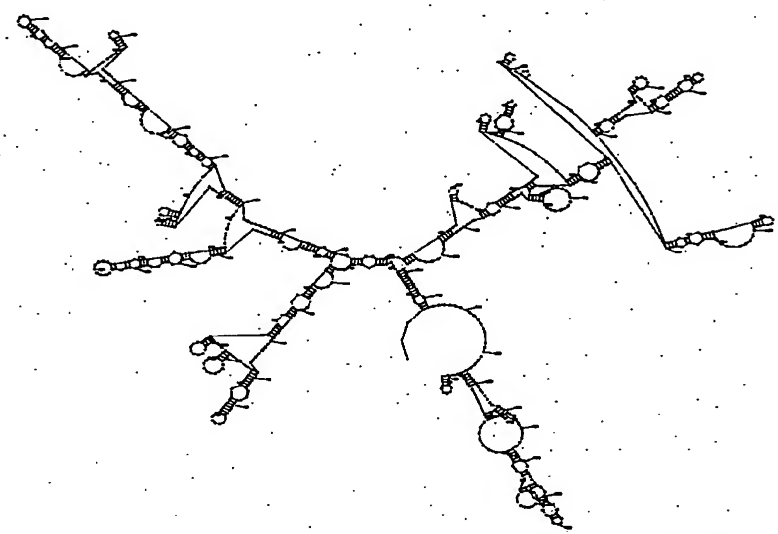
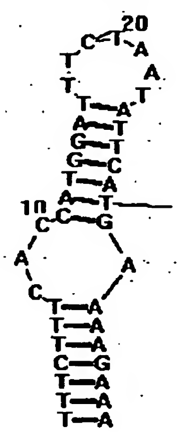
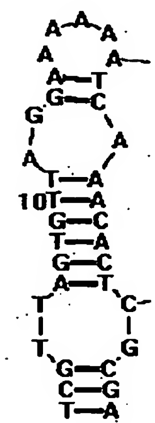


Figure 1. Secondary structure prediction of *B. anthracis* Pag gene 541 – 1560.

As is evidenced by Figure 1, the "folding" of large sequences of DNA reveals several naturally occurring hairpins. The sequences are then isolated from the full sequence and subjected to second structure prediction. Figures 2 and 3 show structural predictions for two of these excised sequences.



Figur 2. Pag 668 – 706
$E_{predict}$ = -4.4 kcal/mol
nt count = 39



Figur 3. Pag 1209 - 1241
$E_{predict}$ = -4.7 kcal/mol
nt c unt = 34

These natural hairpins both appear to be good candidates for use as a molecular beacon, because each contains between about 30 to about 40 nucleotides long and each has a $E_{predict}$ between about -4 kcal/mol and about -12 kcal/mol.

Having confirmed that the selected hairpin(s) satisfy initial selected criteria, a final structural prediction of the sequence in duplex with its complement was computed (Figures 4 and 5). This last prediction was done primarily to ensure that the hybridization of the two DNA sequences, and thus the disruption of the hairpin will be an energetically favorable process. Each of these duplexes have a predicted E value that is about nine to ten-fold greater than the predicted E value for the hairpin alone, and therefore they are expected to favorably form a duplex with their targets.

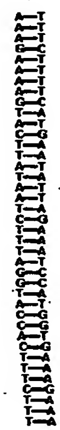


Figure 4. Pag 668 – 706 duplex
$E_{predict}$ = -43.2 kcal/mol

Figure 5. Pag 1209 – 1241 duplex
$E_{predict}$ = -42.6 kcal/mol

The specificity of the hairpin of Figure 2 for its target was supported by a BLAST search of the GenBank database using the Pag 668–704 sequence. The results of this BLAST search are shown below in Figure 6 below. In particular, the BLAST results indicate that only sequences from *Bacillus anthracis*, the target organism, have high scores; whereas other "matching sequences from non-target organisms have significantly lower scores. In this instance, a clear demarcation exists between target scores (of 78) and non-target scores (of 42 and lower). This demonstrates that this hairpin will be specific for its target.

```
                                                                 Score    E
Sequences producing significant alignments:                      (bits) Value

gi|20520075|gb|AE011190.1|  Bacillus anthracis str. A2012 pl...   78   7e-13
gi|16031494|emb|AJ413937.1|BAN413937  Bacillus anthracis par...   78   7e-13
gi|16031492|emb|AJ413936.1|BAN413936  Bacillus anthracis par...   78   7e-13
gi|9280532|gb|AF268967.1|AF268967  Bacillus anthracis plasmi...   78   7e-13
gi|4894216|gb|AF065404.1|  Bacillus anthracis virulence plas...   78   7e-13
gi|10880952|gb|AF306783.1|  Bacillus anthracis isolate BA102...   78   7e-13
gi|10880950|gb|AF306782.1|  Bacillus anthracis plasmid pX01 ...   78   7e-13
gi|10880948|gb|AF306781.1|  Bacillus anthracis isolate 33 pr...   78   7e-13
gi|10880946|gb|AF306780.1|  Bacillus anthracis isolate BA103...   78   7e-13
gi|10880944|gb|AF306779.1|  Bacillus anthracis isolate 28 pr...   78   7e-13
gi|10880942|gb|AF306778.1|  Bacillus anthracis plasmid pX01 ...   78   7e-13
gi|143280|gb|M22589.1|BACPAG  Bacillus anthracis cryptic pro...   78   7e-13
gi|18308294|gb|AC104301.2|  Homo sapiens chromosome 3 clone ...   42   0.038
gi|19033969|gb|AC069286.7|  Homo sapiens BAC clone RP11-261N...   40   0.15
gi|34849950|gb|AC107065.5|  Bos taurus clone rp42-513g13, co...   40   0.15
gi|30962756|gb|AC137820.11|  Medicago truncatula clone mth2-...   38   0.60
gi|30522931|gb|AC123948.4|  Mus musculus chromosome 10 clone...   38   0.60
gi|22552809|emb|AL671857.16|  Mouse DNA sequence from clone ...   38   0.60
gi|11414543|emb|AL355352.16|  Human DNA sequence from clone ...   38   0.60
gi|7768715|dbj|AP001713.1|  Homo sapiens genomic DNA, chromo...   38   0.60
gi|4827077|dbj|AP000178.1|  Homo sapiens genomic DNA, chromo...   38   0.60
gi|4835695|dbj|AP000266.1|  Homo sapiens genomic DNA, chromo...   38   0.60
gi|3132344|dbj|AP000034.1|  Homo sapiens genomic DNA, chromo...   38   0.60
gi|4730836|dbj|AP000102.1|  Homo sapiens genomic DNA of 21q2...   38   0.60
gi|3947430|gb|AC003090.1|  Homo sapiens BAC clone CTA-241I2 ...   36   2.4
```

Figure 6. BLAST® sequence alignment of *B. anthracis* Pag 668 – 704.

<u>Example 2</u> - Hairpins Targeted to *Staphylococcus aureus* Genome

Two DNA hairpins, AH2 and BH2 were designed to incorporate portions of the *Staphylococcus aureus* genome (Genbank Accession AP003131, which is hereby incorporated by reference in its entirety). The AH2 sequence appears to target an intergenic region between ORFID:SA0529 and ORFID:SA0530, and the BH2 sequence appears to target an intergenic region between ORFID:SA0529 and ORFID:SA0530 but including several bases within the latter open reading frame.

A segment of the complete *Staphylococcus aureus* genome was obtained from the GenBank database and the secondary structure of the obtained segment was predicted using computer program RNAStructure version 3.7 (Mathews et al., *J. Mol. Biol.* 288:911-940 (1999), which is hereby incorporated by reference in its entirety). From this predicted structure, two naturally occurring hairpins were identified, one corresponding to AH2 and the other corresponding to BH2.

Having identified these two sequences, these sequences were isolated from the larger sequence and subjected to a second structure prediction as described above. The predicted structure of AH2 is characterized by a predicted free energy value of about -6.1 kcal/mol and the predicted structure of BH2 is characterized by a predicted free energy value of about -3.5 kcal/mol. Both are within the size range of about 30-40 nucleotides.

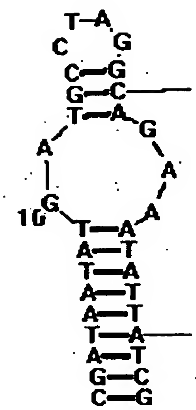


Figure 7
AH2
(E = -6.1)
nt = 33

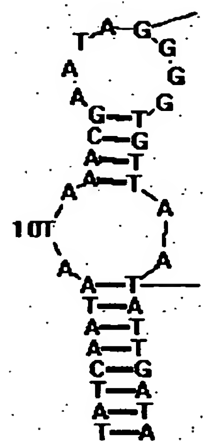


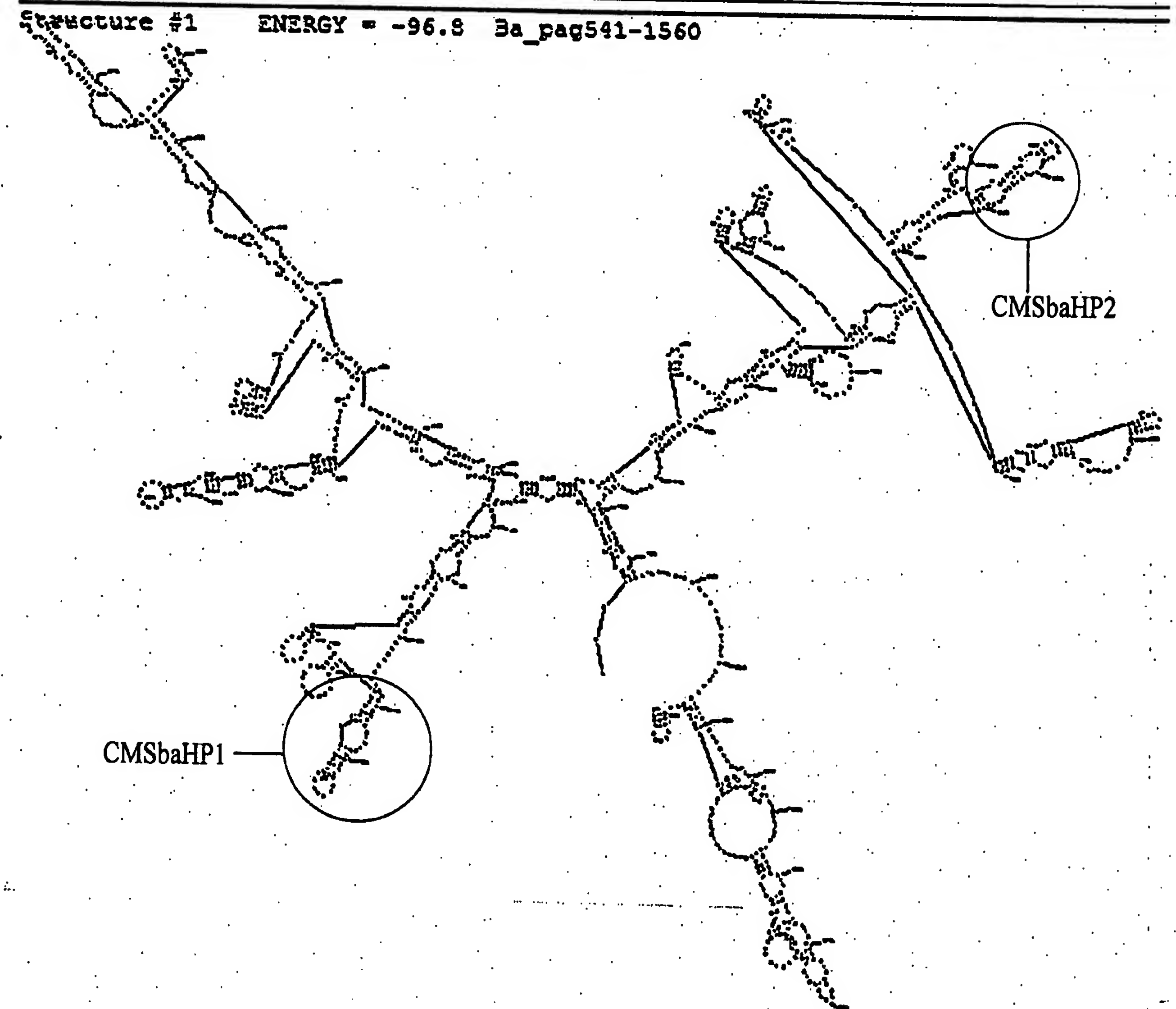
Figure 8
BH2
(E = -3.5 kcal/mol)
nt = 37

Having selected AH2 and BH2, a final structural prediction of the duplexes (AH2 and BH2 with their respective complements) was carried out to determine their predicted E value. The duplex containing AH2 was predicted to have a free energy value of -38.3 kcal/mol and the duplex containing BH2 was predicted to have a free energy value of -39.0 kcal/mol. These values indicate that the hybridization between the hairpin and its target will be an energetically favorable process. A BLAST search was independently performed using the AH2 and BH2 sequences, the results indicating that only segments of the *Staphylococcus aureus* genome contain highly related nucleotide sequences.

This process described above and exemplified in Examples 1-2 has also been performed using *Exophiala dermatitidis* 18S ribosomal RNA gene sequences to identify hairpin probes that can be used to identify the target gene (and organism); *Trichophyton tonsurans* strain 18S ribosomal RNA gene sequences to identify hairpin probes that can be used to identify the target gene (and organism); and *Bacillus cereus* genomic DNA to identify hairpin probes that can be used to identify the target DNA (and organism).

Although preferred embodiments have been depicted and described in detail herein, it will be apparent to those skilled in the relevant art that various modifications, additions, substitutions, and the like can be made without departing from the spirit of the invention and these are therefore considered to be within the scope of the invention as defined in the claims which follow.

B. anthracis Pag gene 541 – 1560 structure number 1



What is Claimed:

1. A method of identifying hairpin nucleic acid probes, the method comprising:

providing a target nucleic acid sequence that is larger than about 100 nucleotides in length;

predicting a folded structure of the target nucleic acid sequence;

identifying the nucleotide sequence of a hairpin within the folded structure of the target nucleic acid sequence; and

predicting a folded structure of the nucleotide sequence of hairpin, in the absence of other nucleotides of the target nucleic acid sequence, wherein the folded structure of the hairpin has a predicted E value of at most about – 3 kcal/mol.

2. The method according to claim 1 wherein the nucleotide sequence of the hairpin is between about 12 and about 60 nucleotides in length.

3. The method according to claim 1 wherein the folded structure of the hairpin has a predicted E value of between about – 4 kcal/mol and about – 12 kcal/mol.

4. The method according to claim 1 further comprising:

predicting a folded structure of a duplex formed between the hairpin and its complement.

5. The method according to claim 4 further comprising:

determining whether duplex formation is energetically favorable.

6. The method according to claim 1 further comprising:

performing a database search for nucleotide sequences that are similar to the identified nucleotide sequence of the hairpin.

7. The method according to claim 6 further comprising:

determining, from the results of the performed database search, whether a clear demarcation exists between scores for target nucleic acid sequences and scores for non-target nucleic acid sequences.

8. The method according to any one of claims 1-7 further comprising:

synthesizing a nucleic acid molecule corresponding to the nucleotide sequence of the hairpin.

9. An isolated nucleic acid molecule prepared according to the process of claim 8.